Subjectivity (2019) 12:210–227
https://doi.org/10.1057/s41286-019-00073-w

ORIGINAL ARTICLE

Digital audiences and the deconstruction of the collective

Laurence Barry[1] · **Eran Fisher**[2]

Published online: 28 August 2019

Abstract
This paper aims at characterizing the change that occurred in audience conception with the advent of big data technologies. We argue that a good place to analyze this change is in the marketing techniques geared to capturing the characteristics of consumers of contents and goods. Some of these techniques are existing statistical tools applied to new kinds of data, others, like predictive analytics, are radically new. Our contention is that online individual actions are now studied, predicted, and managed in the way macroeconomic parameters were analyzed in the past. By changing the perspective on the individual and the group, these new technologies further transform the manner in which an audience is imagined. The conceptions of modern collectives once defined by top-down, broadly defined demographic categories, are therefore transformed or, rather, deconstructed.

Keywords Imagined audiences · Digital audiences · Big data · Algorithms · Predictive analytics

Introduction

At the beginning of the twentieth century, John Wanamaker, a pioneer in advertising, is reported to have said that "half the money I spend on advertising is wasted; the trouble is I don't know which half." The quote appears today in numerous studies on advertising methods in the digital age (see for instance Siegel 2016, p. 27; Cardon 2015, p. 34; Turow 2012): a century after Wanamaker, so it seems, the industry has found ways to avoid the waste.

This paper aims at characterizing the change that occurred in audience conception with the advent of big data technologies. This research stemmed from an initial project to understand the manner in which digital media curators (such as Facebook and Google) recommend online readers with further readings, and how digital

✉ Laurence Barry
barryl@smile.net.il

[1] Hebrew University of Jerusalem, Jerusalem, Israel

[2] Open University, Ra'anana, Israel

audiences are thus imagined by those curators. It soon became clear however that the techniques at stake do not concern solely readers, but treat more generally the data produced by visitors of any site: online recommendation systems are gears towards proposing further readings or promoting purchases, and there is no conceptual nor technical difference between these two aims. Digital marketing techniques thus contribute to blurring the distinction between readers/spectators (audiences of the traditional kind) and potential online customers (audiences as consumers), and imagine these audiences in the same manner.

Furthermore, algorithmic curation implies the use of new technologies: some are existing statistical techniques applied to new kinds of data, others are radically new. For the sake of this research, we analyzed technical papers in data science applied to marketing that describe how recommendations or predictions of specific online actions are algorithmically produced. Since this literature is an already extensive body of texts, we decided to focus on one specific family of models that was particularly insightful; it takes as an input the online behaviors (e.g., texts from Twitter, blogs, or Facebook likes) and fit the output to adjust to the results of one or several psychological tests taken separately by the users.[1] This methods creates a profile based on the individual's navigation behavior and is reputed to infer his personality better than previous techniques. Without pretending to understand the depth and complexity of these models, we found this literature useful in laying bare both the assumptions made and the conclusions reached by data analysts. In presenting their methods and results, we argue, the authors convey some sense of how they see the audiences they are trying to target, a sense we have aimed at outlining here. Our contention is that online individual actions are now studied, predicted, and managed in the way macroeconomic parameters were analyzed in the past.

Finally, these new technologies and the perspective they adopt towards the individual and the group transform the manner in which an audience is imagined. Building on the new knowledge, current marketing literature indeed emphatically insists on the possibility to adjust and personalize recommendations of products, contents, etc.… *to the specific individual*. The combination of the technical shift and the new marketing discourses creates a new imaginary of audiences; digital audiences are thus conceptualized with data and algorithms (Goriunova 2019a), by data analysts and marketers. Furthermore, we argue that this trend is symptomatic of a larger phenomenon, where the conceptions of modern collectives once defined by top-down, broadly defined demographic categories, are being transformed or, rather, deconstructed.

The emergence of the digital audience

The term "audience" has a long history, one that started maybe with the ceremony of "an audience with the King" (Livingstone 2005, p. 29), as an encounter between the king and his subject, where the speaker and his royal audience were identifiable

[1] See Reference section for the corpus.

to one another. In modern times, by contrast, the audience comes to define "a large number of *unidentifiable people*, usually united by their participation in media use" (Hartley 2002, p. 11, emphasis added). As the totality of readers, listeners, or spectators of a given media, it is an aggregate, the nature of which was for long inferred from its published content.

The modern audience has thus an abstract shape; it is an imagined community, where no specific member is readily reachable by or known to the speaker (Anderson 2006, pp. 6–7). Following Habermas, one can say that it retains some characteristics of the Kantian public: for Kant, addressing the "public" meant indeed going out of one's private, concrete identity in order to speak to "a society of world citizens" (Habermas 1991, p. 106). Yet while Kant's public was imagined as universal, most imagined communities are more restricted; for Anderson, the imagination of a nation also entailed the definition of its limits. At the end of the nineteenth century, the census was thus supposed to allow an exact circumscription of the community at stake: "the fiction of the census is that everyone is in it and that everyone has one, and only one, extremely clear place" (Anderson 2006, p. 166).

The census could classify people within (arbitrarily defined) categories based on their own answers. But the broad media audiences of the twentieth century are more difficult to grasp (Turow and Draper 2014, p. 644): addressees are mainly anonymous, and yet there is an attempt to speak to "everyone-as-someone" (Scannell 2000). Our assumption here is that content producers need to imagine to whom they are speaking (Ang 1991; Litt 2012; Matthews 2008; Marwick and Boyd 2011), and that this imagining is conditioned by the medium of communication (Cover 2006; Livingstone 2003; Napoli 2010) and the technological tools at their disposal. Mass media audiences were thus imagined as deriving from the message delivered, or the hour and day of the broadcast, a description that remains rather amorphous. Hartley (2002, p. 11) contends that "given the varying demographics of this group, (…) the concept itself is a means by which such an unknowable group can be imagined. Naming an audience usually also involves homogenizing it, ascribing to it certain characteristics, needs, desires and concerns." For Livingstone by contrast, the work of homogenization has intrinsic limits as long as audiences remain envisioned *as aggregates of individuals* with common characteristics but also specific desires (Livingstone 2005, pp. 24–25; See also Anderson 2011; Napoli 2010). Some further argue that the difficulty to produce such an abstraction led some journalists to imagine their audience in the form of their supervisors (Turow and Draper 2014, p. 645; Anderson 2011, p. 553).

The imagining of audiences seems therefore to be caught in a tension between the need for homogenization on the one hand, and the remnant image of an aggregate of specific, mostly unknown individuals on the other: in Dumont's terms, the audience is imagined both as *societas*, a collection of individuals, and as a corporate whole, as *universitas* (Dumont 1983, p. 98). Wanamaker's complaint takes place in this context; advertising in a specific media was for long a blind decision based on a few traits attributed to its audience.

According to Turow, the 1980s marked a shift in these attempts: the essentialized conception of the audience (where its characteristics were conceived as an objective yet unknown identity) turned into a constructivist approach (Turow and Draper

2014, pp. 646–647). This shift in conception was enhanced by the Internet in the 1990s and the increasing power of private institutions to shape its use. For Turow and Draper (2014, p. 649), despite some optimistic vision of the internet as a space of empowerment for individuals that would freely choose what they see, read, and consume online, it is actually dominated by “interconnected corporations.” Internet use and audiences are actually shaped by “the new media buying system” that developed with big data technologies in advertising (Turow 2012). This paper comes to shed light on the technical aspects of this shift and the implications for the conception of audiences by the media professionals who target specific individuals or groups in the digital age.

In their practice of targeting, these professionals contribute to the creation of new objects of knowledge, digital audiences, and “digital subjects” (Goriunova 2019a). We are careful though to make a distinction between the imagining of audiences, as groups or individuals, and the processes of subjectivation they imply. Subjectivity is not wholly at the hands of structural forces, and socio-technical apparati cannot be said to have a deterministic and uni-directional effect. As Giraud points out “subjectivity <is> being constructed through engagements with technological infrastructures, but in ways that do not seamlessly commodify subjectivity and instead create friction and space for resistance” (Giraud 2015, p. 125).

The shift happened, we want to argue, in two stages. Following Mayer-Schönberger and Cukier (2013, p. 82), one can distinguish between the phases of digitization and datafication that accompanied the rise of the digital age. The digitization phase consists in “turning analog information into computer format” (Mayer-Schönberger and Cukier 2013, p. 83). Digitization isomorphically creates digital objects that give an online representation of their analog referent. Most offline companies thus developed a site for selling online their retail products. But as a consequence, they turned into “internet publishers” and potential advertisers for third parties (Sternberg 2013). Online retailers thus take part in the abovementioned media-buying system that slowly erased the distinction between content sites and commercial ones (Turow 2012, p. 2744; Pasquale 2015, p. 71): both have now their own audience. Today, the concept of an “audience” includes therefore a large variety of groups, such as online potential, new or recurring buyers: these are “audiences of consumers” (Cheney-Lippold 2011, p. 167). This obviously contributes to the commodification of subjectivity mentioned by Giraud: recommendation systems indeed treat equally data from readers and buyers. Some insist however on the interactive nature of the new internet, showing how audiences are now not solely receivers of content but also senders (Cover 2006; Livingstone 2003; Napoli 2010), thus transforming into “participatory audiences” (Giraud 2015). We would rather focus here on the internet and the web 2.0 as *a marketing tool* (Constantinides et al. 2009), and how this tool is used to imagine audiences.

The datafication phase of the digital age is the one that really changed, we would like to argue, the manner in which audiences are technically conceived. It occurred when the information became “usable not just for human readers but for computers to process and algorithm to analyze” (Mayer-Schönberger and Cukier 2013, p. 83) and, we would like to add, when information was generated from online activity itself. Data thus became organic to human online behavior. In the case of audiences, it means that the concept is itself quantified, and becomes an object of analysis for "data crunchers." In this perspective, the participatory audiences that produce content also feed the models with additional data collected on the users. The abstract imagining is thus doubled or replaced by technical tools to analyze the set of visitors of a site (consumers or readers) and further target "lookalikes." The marketing industry is now replete with technologies that offer to "design and understand one's audience": Outbrain, Acxiom, Connexity, to name a few, all have developed their own algorithms to both understand and enhance the audience of a site in order to increase its sales.[2]

This double evolution significantly transforms, we would like to argue, the practice of imagining the audience. Indeed, traditional marketing consisted in finding "the right place to advertise while holding everything else constant" (Facebook 2017): studies could prove for instance that comedy shows were best suited for food commercials, whereas drama shows were optimal for pain reliever commercials (Harvey). "People-based marketing," launched in 2014 by Facebook, turns the focus towards the users via their data: in a first stage, it aimed at finding "the right *group* of people, while holding everything else constant"; in its most recent development, it allows to identify a customer across his online devices in order to adjust the offer to his specific navigation behavior. The rule of the marketing game, and slogan, is now "marketers don't hold anything constant. *We now can find a person instead of large groups like "Adults 18–34"* (…) and reach them on whatever device or platform they may be on" (Facebook 2017, emphasis added). The audience is thus not *inferred* from the broadcasted content, but aggregated from specific consumers or internet users, on whom is placed the inference. The traditional audience practitioner used to be the producer of content; he is replaced today by the data scientist that gathers among users the best prospects in order to create segments adjusted to a given target, the others being considered as "waste" (Turow 2012, p. 1875). In Wanamaker's terms, the data scientist cherry picks the relevant half that is now considered as the audience. The content has thus become secondary to this data analysis (Turow 2012, p. 2595). As Pasquale puts it,

> The ad buyers argued that it's not space on paper or pixels on a website that matters to them, but audiences; that's what they were looking to buy. In other words, the context of the advertisement was mere background: what really mattered was *data on who was looking at the content* (Pasquale 2015, p. 96, emphasis added).

With the Internet, the amorphic imagining of audiences is thus replaced by the analysis of the data collected. This goes hand in hand with an increased granularity of the data at stake. Both the census communities and the media audiences were

[2] Outbrain's slogan for businesses and brands is thus "reach new audiences" (see https://www.outbrain.com/amplify/, accessed 12/12/2017); Acxiom proposes to "to create ideal audiences through advanced segmentation, scoring and modeling strategies" (https://www.acxiom.com/how-we-can-help/unify-offline-and-digital-data/ accessed 13/12/2017); Connexity offers a solution to grow one's "seed core audience" (Connexity 2016, p. 1);.

imagined on a limited set of information, mainly demographic parameters, derived from questionnaires. The big data era is characterized by the combination of the capacity to collect freely a huge amount of data on each specific user, with the technology to analyze them (Andrejevic and Gates 2014, p. 186). For Turow, one of the major devices that allowed the shift was the gathering of internet traces thanks to cookies (Turow 2012, p. 1029); others mention the generalization of smartphones in the last decade (Mayer-Schönberger and Cukier 2013, p. 90; Lambiotte and Kosinski 2014, p. 1937; Weed 2017). We will argue in the next part that the traces left by individuals in social media (posts, tweets and likes), combined with the techniques to make sense of these data, take the audience imagining to a new level. Demographic parameters are now perceived as insufficient to describing the users. Some speak of micro-segmentation that consists in capturing "both demographic and behavioral data" (Teich 2016) or of a "hyper-segmentation revolution" (Weed 2017; see also Turow et al. 2015, p. 468); others even denounce the stereotyping implied by using demographics altogether (Brown 2016).

The ultimate goal of the search for increasingly granular data is being able "to *recognize* a customer or prospect, on all their digital devices, and on any marketing platform" (Acxiom 2016; see also Quancast 2016, p. 11). In this perspective, Weed suggests in his recommendations to Google to move from mass-marketing to mass-customization, "from focusing on averages to individuals," thus creating "segments of one" (Weed 2017). Various techniques are available today to help create such singular segments. They all start by producing a data-driven understanding of the *individual* consumer that implies, we would like to argue, a significant transformation of the imagining of audiences. How this is done in practice is described in the following part.

Predictive analytics: from averages to individuals

The abovementioned injunction to move from averages to individuals in marketing relies on the belief in the capacities of predictive analytics. We argue in this part that these techniques blur the distinction between disciplinary and security mechanisms as introduced by Michel Foucault (Foucault 1995, 2009). Following Ewald (2011), our contention is that a new conceptual relationship between the individual and the group is taking shape with the advent of big data; a new technique for the apprehension of the collective is emerging which transforms its imagining, as will be described below.

From individuals to aggregates…

In nineteenth century disciplinary regime, Foucault claims, individual behavior was observed and measured against a desired norm. In the ideal prison of the *panopticon* for instance, each prisoner is individualized in a specific cell where he is continuously scrutinized and registered. Discipline observes, but also corrects: thanks to continuous examination, ranking, and sanction, it seeks to bring each and every one closer to the desired behavior, posed as normal (Foucault 1995, pp. 177–183). This disciplinary approach thus tends to imagine the group in the form of this singular preferred behavior, the norm imposed to all.

The security mechanisms by contrast do not try to physically correct behaviors, but focus instead on this "other level of reality" that appears with statistical measurements: the regularity of death rates, morbidity incidence, accidents, and the likes suggests the existence of a collective subject, the population (Foucault 2009, p. 104). The population is then imagined through these steady averages that prove its existence. This shift marks for Foucault a novel approach to the collective: instead of a singular behavior that would represent the whole, the group is understood as having a distribution of behaviors, the vast majority of them being considered normal (Foucault 2009, pp. 57–58).

The individuals that feed the population are not the object of knowledge anymore: by contrast with the disciplinary techniques, the security mechanisms aim at securing the population as a whole and admit that the individual remains unknown (Desrosières 1988, p. 46; Desrosières 1998, p. 55; Foucault 2009, pp. 9–10). From this viewpoint, the regulation of the collective entailed the renunciation of the disciplinary attempt to know exhaustively and act on each single body. Besides, although both techniques of government function simultaneously in modern societies, their focus on different objects of knowledge makes them orthogonal to one another: the knowledge of the population requires becoming blind to the individual, and vice versa.

Yet the statistical approach needs personal information to build upon, and therefore the security mechanisms are anchored on the disciplinary ones. More precisely, if one defines discipline as the combination of surveillance and punishment, the security regime maintains surveillance, now combined with the statistical treatment of the data thus gathered. This statistical approach further implies a series of human interventions to organize the information in a generalizable manner. Indeed, for most of modernity, the data were not readily available and had to be gathered through questionnaires or archives. It always resulted from some arbitrary decisions necessary for the codifying of reality, thus reducing its diversity to a limited series of variables; those helped build the "contact zone" that could then be analyzed (Desrosières 1998, p. 105). The statistical approach, prior to the Internet, thus implied a quantification of reality, different from both the digitization evoked in the previous part and from the examination reports and ranking systems created for disciplinary purposes.

This quantification served the security mechanisms in the construction and the management of the collective: it created variables to be interpreted at the aggregate level. As Desrosières puts it, the data were collected horizontally each line representing the data of one person, but the statistics created vertical meaning. Each column, or variable, was made of the collection of the coded answers of the participants to a specific question, as in the censuses mentioned in the previous part. For most of the twentieth century, statistics and econometrics took the columns as objects of knowledge to measure correlations and look for causation, in order to give conclusions expressed as macro-phenomena (Desrosières 2008, p. 140, 2014, p. 169). While the models became more complex over time, the general approach remained vertical.

… and vice versa

The real novelty brought by predictive analytics, we want to argue, is that they lead to a transposition of this approach: instead of looking at the vertical, aggregate meaning of *a variable*, the data are analyzed as *features* of the specific observation. Siegel characterizes the distinction between traditional statistical methods called forecasting, and predictive analytics as follows: "whereas forecasting estimates the total number of ice cream cones to be purchased next month in Nebraska, predictive analytics tells you which individual Nebraskans are most likely to be seen with cone in hand" (Siegel 2016, p. 16). The analysis remains focused on the typical attributes of the individual at stake. We will try to show in this section the outlines of the method on a specific family of models: they are all focused on the analysis of online texts for the characterization of the personality of users, further taken as input for targeted marketing.

Text analysis for the sake of personality characterization is not new (Goldberg 1993). Earlier models, developed in the 1990s, implied the creation of categories of words in order to measure the frequency of their use in a given text. These categories were the explaining variables and included both grammatical and semantical indications. In analogy with the census or other questionnaires, the elaboration of these variables relied on manual codification done by experts. Thousands of words were examined and attributed to pre-defined categories, thus reducing the diversity to a few variables (Tausczik and Pennebaker 2010).

More recently, growing computer capacities rendered possible the abandonment of categories in favor of the processing of words themselves, shifting the domain to big data analyses (Yarkoni 2010; Schwartz et al. 2013; Arnoux et al. 2017). This shift has a series of implications. First, the number of variables becomes obviously tremendous, moving from a few dozens to thousands. Yet with their multiplication, the information each variable conveys becomes minimal. Indeed, the questionnaires of the security era were complete, in the sense that there were no missing points or questions left unanswered. By contrast, one of the difficulties faced by data scientists is that the features of big data are sparse: the click on a given page or the use of a specific word is rare by nature (Chen 2009, p. 211; Li et al. 2010; Schwartz HA et al. 2013, p. 2). As a consequence, the contribution to meaning of each feature on the aggregate is low: the use of these features hence forces, we argue, the abandonment of the macro-level where insights were drawn from correlations between variables.

Second, the absence of a priori codification is often interpreted by data scientists as a warrant of being closer to the truth of the phenomena to analyze.[3] This goes hand in hand with their claim that the dataset, because of its size, would have become exhaustive and truly objective (Mayer-Schönberger and Cukier 2013, pp. 19–32; Boyd and Crawford 2012, p. 666). In the case of online posts, data scientists indeed contend that:

[3] Desrosières mentions the polemics around quantification as a growing issue in the production of statistical results (Desrosières 2014, pp. 33–35), a problem that is apparently lifted with big data.

> Online social media such as Facebook are a particularly promising resource for the study of people, as 'status' updates are self-descriptive, personal, and have emotional content. *Language use is objective and quantifiable behavioral data, and unlike surveys and questionnaires, Facebook language allows researchers to observe individuals as they freely present themselves in their own words*. (Schwartz et al. 2013, p. 13, emphasis added).

The family of models studied here are usually built following the same method; in a first stage, a group of people, constituting the "training set," both accepts to fill a personality questionnaire and allows access to their online profile, either as posts on Facebook, tweets, or longer blogs. The questionnaire is used to establish a personality profile, along traditional models (usually the "big five" model, see Goldberg 1993; Brown 2016; Chen et al. 2015). The algorithms then take as input these online texts, and aim at predicting the personality traits as inferred by the questionnaire, taken as "ground truth," the reality to be modeled (Lambiotte and Kosinski 2014; Schwartz et al. 2013; Yarkoni 2010). Once valid, it allows to infer the personality of *any* user outside the training set, without the need to have him take a personality test. The output is usually a five dimensions score (following the "big five" model), which informs the marketing recommendations.

These algorithms thus actualize a hybrid combination. Their aim being to trigger the urge to look at the next offer or recommendation (Karakayali et al. 2018), they undoubtedly serve a government exercised through affects, suggestions, and the bypassing of "conscious awareness or control" (Blackman 2019). They thus assume a subject that is imminently suggestible. Yet they do so while unreflectively relying on psychological models developed under the assumption of an essentialized personhood or "an essentialist inner mechanics of psychological functioning," (Blackman et al. 2008, 10).[4] Historically, this strand of psychology, known as "the science of personality" (Koopman 2019, p. 66), takes the "normal" subject to be rational and self-contained, and associates suggestibility with "a lack of a set of competences that would enable the subject to withstand social influence and therefore separate themselves from others"; in other words, in this strand of psychology, suggestibility is nothing other than "a left-over animality" (Blackman 2008, p. 34), or what precisely *cannot* be grasped via the models.[5]

While the personality questionnaire serves as a touchstone, the claimed result of these analyses is that the algorithmic inference of a user's personality from the texts posted online actually gives better results; in some cases, the algorithm is even compared to close friends and families knowledge of the user and is shown to perform better (Youyou et al. 2015). The algorithm is thus taken to be a better judge of personality than both the experts that created the questionnaires and the close

[4] The "big five" model is indeed inherited from the Cambridge psychometrics laboratory founded by Galton and Cattell at the end of the nineteenth century (Goldberg 1993; see also Koopman 2019, pp. 66–107). As Koopman (2019, p. 67) suggests "although it is easy to cast doubt on the accuracy of these instruments, it is really their functional success that should impress us."

[5] Their reappropriation within a datafied and algorithmic context obviously displaces their meaning and implications. The analysis of this displacement remains however out of the scope of this paper.

surroundings of the individual, showing the double distanciation from security and disciplinary techniques, as will be detailed below.

First, the potential bypassing of questionnaires indeed highlights the shift from the security techniques and its global perspective. The questions in surveys were necessarily homogenizing, creating groups of people that were assumed to have similar characteristics; as Desrosières puts it, the categories created “classes of equivalence” (Desrosières 1998, pp. 1–12). With online data gathered directly from personal profiles, the information is taken to be closer to the true individual behavior as “expressed freely,” in the absence of any constraints imposed by the guidance of the questionnaires. The first consequence, we argue, is that each individual remains irreducibly different from the others. Reflecting on this point, Berns and Rouvroy contend that:

> The profile “linked” to an individual’s behaviour could itself be tailored perfectly efficiently (…) to the extent of it seeming *as though all discriminatory categories are avoided*, and even of being able *to take into account what is most specific to each individual*, what is most distant from big numbers and averages (Berns and Rouvroy 2013, p. 19, emphasis added).

The information is also considered as more truthful since it escapes arbitrary categorization that aggregates users. Trying to predict depression based on online posts, De Choudhury et al. symptomatically mention that previous analyses were necessarily relying on “high-level summaries” (De Choudhury et al. 2013, p. 129). The fragmentation is thus taken to reflect the movement of life itself, in a manner that cannot be grasped by the human mind. Some data scientists hence claim that:

> First, computers have the capacity to store a tremendous amount of information, which is difficult for humans to retain and access. Second, the way computers use information—through statistical modeling—generates consistent algorithms that optimize the judgmental accuracy, whereas humans are affected by various motivational biases (…) In the future, people might abandon their own psychological judgments and rely on computers when making important life decisions, such as choosing activities, career paths, or even romantic partners. It is possible that such data-driven decisions will improve people’s lives (Youyou et al. 2015, p. 1039).

Second, although centered on the individual, predictive analytics also differ from disciplinary techniques since they apply digital and statistical tools rather than physical constraints. In order to adjust as much as possible to the individual, the models are indeed built on behavioral data: “ad clicks and views, page views, search queries and clicks” (Chen et al. 2009) are all considered the same. More generally, the input data are of two main kinds, data and meta-data:

> The data consists of information actively and often voluntarily entered by the user, such as personal details, photos, comments, messages, search terms, bids in auctions, payment information, and connections with “friends”, as well as passive footprints such as the duration they spent on the website, what pages were browsed, in what sequence, the website that referred them to the web-

> site, the Internet browser and operating system used, location, and IP address (Shmueli 2017, p. 2).

Yet the distinction is irrelevant, as far as the algorithm is concerned: both are treated with the same techniques and convey the same kind of computerized information. As concerns our models focused on text analyses, it is not the meaning of what is said that conveys information, but the *use* of specific words as *behavior*. For the sake of its computerized interpretation, language is indeed equivalent to "*quantifiable behavioral data*" (Schwartz et al. 2013, p. 13, emphasis added). As Zuboff puts it: "Google is 'formally indifferent' to what its users *say or do*, as long as they say it and do it in ways that Google can capture and convert into data" (Zuboff 2015, p. 79; see also Andrejevic 2016, p. 36).

Furthermore, the tremendous amount of data collected renders possible the *statistical* treatment of the individual. For Andrejevic this means that "drone logic only individualizes through the lens of the aggregate" (Andrejevic 2016, p. 35). We would like to argue instead that the individualization is paradoxically obtained via *statistics*. In the security era, statistics were used to create the aggregates; they now serve the individuation of statistical selves. Compared to the data collected through questionnaires, the table has indeed become as large as it is long; in other words, there are often more variables than users (Kosinski et al. 2016, p. 496). Hence, the recommendation models treat indifferently the lines and columns of the matrix of items as ranked by users. The symmetry is often expressed in the data science literature as a pair between an individual and an item, an ad or a like (Leskovec et al. 2014, p. 308; Kosinski et al. 2013; Kosinski et al. 2016, p. 494).

Finally, the basic technique for validating the prediction consists in cross-validation, i.e., the splitting of the dataset into two groups of users: one being the training set that is used to calibrate the parameters, the other being the test set. The latter consists of another group of users, on whom the accuracy of the prediction is being measured (Kosinski et al. 2016, p. 502). The end result of the study is not a formula expressing a relation between variables, but a score attributed to individuals: hence, the aggregate is never more than the collection of individuals that constitute it, with a predictive score attributed to each. In Desrosières's perspective, the approach remains horizontal all along.

Imagined digital audiences?

It is now possible to consider how these new algorithms affect the concept of audience and, more generally, the manner in which collectives are being imagined in the digital age. According to Bauman and Lyon, people continue today to "get clustered into crude population segments so that marketers can treat them differently depending on their consumer behavior" (Bauman and Lyon 2013, p. 61). Our argument here will be that, stemming from the techniques of prediction focused on the individual, the imagining of audiences is done bottom-up rather than top-down as was the case in the past. Hence if groups are still used for marketing purposes, the knowledge of these groups and the imagination it nourishes have significantly changed.

It is often said that the digital categories of the algorithms create identities that keep no connection with their real-life counterparts: age, gender, or “being a terrorist” are redefined based on one’s online behavior. Cheney-Lippold therefore uses quotation marks to distinguish between online ‘categories’ and offline categories (Cheney-Lippold 2017, p. 10). Online ‘categories’ are characterized by their fluid pattern. As Cheney-Lippold puts it: “the gender of the same user might change from male to female if enough user data, such as the addition of certain web sites that user visited, are presented to statistically identify that user with a different gender” (Cheney-Lippold 2011, p. 168). This point is confirmed by marketers that insist on the importance of recency: “fresh audiences and recent data points ensure that your ads are reaching new consumers most receptive to your messaging” (Connexity 2016, pp. 2–3). The “novelty” is here defined by the changing receptiveness of existing persons rather than new persons altogether, the aim being to catch “temporal behavior” (IBM 2017; see also Lee et al. 2014).

Therefore imagined digital audiences build on “digital subjects” (Goriunova 2019a) or “alien subjects of Artificial Intelligence” (Parisi 2019; see also Parisi 2016) as they are “formed in the computational ensembles out of data, models, and various other analytical operations” (Goriunova 2019a, p. 1). To be sure, they try to bare one-to-one relations with human individuals, but they are wholly imagined and comprised within the context of a digital universe.

In Cheney-Lippold’s view, the algorithms actually create categories, albeit of a new kind: “algorithms allow a shift to a more flexible and functional definition of the category, one that de-essentializes gender from its corporeal and societal forms and determinations while it also re-essentializes gender as a statistically-related, largely market research-driven category” (Cheney-Lippold 2011, p. 170). Our contention is that the statistical definition of gender is not a category but a number in a continuum that thus escapes the assumption of equivalence and homogeneity always implied by categories (Desrosières 2008, p. 122). Even less socially ambiguous constructs than gender can be undermined by this outlook. As Goriunova (2019b) suggests, even facial recognition softwares, which are supposed to create an unproblematic link with a biological reality, are in fact susceptible to such abstraction.

In a recent study on linguistic styles online, Bamman et al. (2014, p. 135) hence question the use of demographic variables such as gender, age, race, or geographical origins: for them, behind these attributes, “there is often an implicit assumption that linguistic choices are associated with immutable and essential categories of people.” Specifically working on gender, they contend that the data do not support a dichotomous split that oversimplifies reality. They conclude with the following general insight:

> While the statistical relationships between word frequencies and gender categories are real, they are but *one corner of a much larger space of possible results* that might have been obtained had we started with a different set of assumptions (…) Machine learning offers a bountiful harvest of modeling techniques *that minimize the need for categorical assumptions* (…) we hope that a more nuanced model might allow statistical reasoning *on the level of individual micro-interactions* (Bamman et al. 2014, p. 153, emphasis added)

They thus recommend the discarding of a priori categorical assumptions altogether, because those mislead researchers: observations that do not fit into the categorical distribution of behaviors are filtered out as outliers, while they actually contain insightful information.

In the first step of the study described in the previous part, essential, humanly pre-defined, categories (the five big traits of personality) were shown to be inferable from online behavioral data. Bamman et al. illustrate the fact that the algorithm actually functions without categories; those can be reproduced for the convenience of marketers, but the "crude categories" evoked by Bauman and Lyon are in fact unnecessary. This further explains their fluid definition mentioned by Cheney-Lippold; categories are simply not part of the structure of the data. As LeCun et al. put it by reflecting on their own work: "the key aspect of deep learning is that <the> layers of features are not designed by human engineers: they are learned from data" (LeCun et al. 2015, p. 436).

Technically, the practitioner can summarize his audience in one number, the predicted probability of response, or score: the delimitation of the group results from an arbitrage between efficiency, on the one hand, that gives preference to high score levels but implies a small audience size, and volume necessities on the other that lead to its increase despite the loss in overall efficiency (Siegel 2016, p. 14; Chen et al. 2009, p. 210). The group thus defined is not in any way describable by demographics but simply corresponds to the best possible group in predictive terms, given a desired size (Cardon 2015, p. 53). Besides, the complexity of the models at stake makes it impossible to give any explanation either to the contours of the group or the score of a specific individual (Pasquale 2015, pp. 4, 24).

The modern imagining of the collective in the broad demographic terms described in the first part is thus bypassed by the computerized capacity to produce groups, the homogeneity of which resides only in a common score level. This seems to realize Weed's wish of "segments of one" (Weed 2017): each user is and remains specific. Recommendation systems are now apt to choose among a vast majority of possible offers (content recommendation or other) the one best adjusted to a specific user, defined by his online behavior (Cheng et al. 2016; Zeldes et al. 2017). Facebook's people-based marketing is becoming a reality: the consumers, disassembled into the features of their navigation, are offered products that fit their online behavior, rather than being categorized as belonging to a specific audience.

One of the methods for fitting an item to an individual relies on the isolation of some regularities that characterize one user as similar to other users taken from the training set (Leskovec et al. 2014, p. 30). The regularities or patterns evidenced by the algorithms are determined at the individual level. Interestingly, the government of population at the dawn of statistics also relied on the discovery, this time by human calculations, of regularities. But those were very different from the regularities pinpointed by the algorithm; they were possible only at the aggregate level of statistics. This collective imagined level is deconstructed, we would like to suggest, by predictive analytics. Focused on digital behaviors, these techniques rely on statistics and mathematics in order to predict future individual actions rather than macroeconomic results. The collective is thus nothing other than the collection of individually known singular elements.

For Zuboff, this constitutes a major departure from the neoliberal order. Indeed, "the unknowability of the universe of market transactions" was at the heart of Hayek's neoliberal doctrine (Zuboff 2015, p. 78). The algorithmic knowledge of online transactions leads to a reversal; the market is not apprehended top-down as "ineffable and unknowable"; it is rather constructed bottom-up by the micro-processing of each and every transaction.

Drawing on current capacities to collect and analyze vast amounts of singular transactions, Pentland further argues that a new social science, social physics, is taking shape, in which the collectivity is understood as the micro-relationships of individual nodal points:

> Social phenomena are really made up *of billions of small transactions between individuals*—people trading not only goods and money but also information, ideas, or just gossip. There are patterns in those individual transactions that drive phenomena such as financial crashes and Arab springs. We need to understand these micropatterns *because they don't just average out to the classical way of understanding society*. Big data give us a chance *to view society in all its complexity, through the millions of networks of person-to-person exchanges* (Pentland 2014, pp. 10–11, emphasis added).

Big data and predictive analytics are thus not merely allowing old ways of seeing the public on a massive scale but provide "new ways of controlling how publics come to be represented and so understood" (Kennedy and Moss 2015, p. 1), or in our terms, a new epistemology.

Conclusion

In his essays on modern individualism, Dumont claimed that a tension continued to exist between a vision of society as a collection of individuals and a more traditional conception of the group as a homogenized whole. In his view, the social sciences continued to perpetuate as a scientific discipline, the consciousness of the social whole.

The tension between the whole and its parts was visible in early marketing practices; they relied on surveys for the characterization of audiences in a process of categorization and homogenization, similar to the one performed by censuses and other statistical inquiries. With big data technologies and the datafication of the world, the arbitrary quantification of phenomena is now obsolete; essentialized categories are not necessary any more for the treatment of information. Furthermore, the accumulation of data at the individual level allows the application of statistical techniques, once focused on the collective, at this same individual level: statistical techniques are now applied to the individual. Big data technologies have thus rendered possible the treatment of populations without the need to rely on aggregate indicators such as averages.

The imagining of audiences and collectives is therefore significantly transformed. While prior marketing analyses estimated an overall level of performance, predictive analytics aim at cherry picking specific individuals as part of the audience thus built. The audience is not constituted of people recognized as similar in any specific demographic characteristics any more, but simply as fulfilling a certain level of predicted score. The "people-based marketing" advocated by Facebook and now adopted by many marketers entails a further abandonment of the collective level, as it offers to adjust to each navigation profile its optimized advertising.

The traditional demographic description remains available, but it is inferred bottom-up, based on more granular available information on online behavior. In the distance, the picture of the group may look the same, but if one decides to get closer, the features of each individual remain accessible. The treatment of the collective is now obtained by the statistical management of individual behaviors. From Dumont's viewpoint, one might wonder whether the *societas*, the collection of individuals, has not finally managed to wipe out the *universitas,* the group as a whole.

Funding This research was funded by Grant Number 696/16 from the Israel Science Foundation.

Corpus

Acxiom. 2016. The State of People-Based Marketing. https://marketing.acxiom.com/rs/982-LRE-196/images/TheStateofPeopleBasedMarketing_2016-10-11_acxiom.pdf. Accessed 1 June 2019.

Arnoux, P.H., A. Xu, N. Boyette, J. Mahmud, R. Akkiraju, and V. Sinha. 2017. 25 Tweets to Know you: A New Model to Predict Personality with Social Media. *AAAI Publications, Eleventh International AAAI Conference on Web and Social Media* 2017: 472–475.

Bamman, D., J. Eisenstein, and T. Schnoebelen. 2014. Gender Identity and Lexical Variation in Social Media. *Journal of Sociolinguistics* 18 (2014): 135–160.

Brown, J. 2016. Omni-Channel? https://www.acxiom.com/blog/omni-channel/.

Chen, J., E. Haber, R. Kang, G. Hsieh, and J. Mahmud. 2015. Making Use of Derived Personality: The Case of Social Media Ad Targeting. *AAAI Publications, Ninth International AAAI Conference on Web and Social Media*, 51–60.

Chen, Y., D. Pavlov, and J.F. Canny. 2009. Large-Scale Behavioral Targeting. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 209–218.

Cheng, H.T., L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anerson, G. Corrado, W. Chai, M. Ispir, R. Anil, Z. Haque, L. Hong, V. Jain, X. Liu, H. Shah. 2016. Wide and Deep Learning for Recommender Systems. arXiv:1606.07792.

Connexity. 2016. Peek Behind the Curtain: The Basics Behind Connexity Audience Modelling. http://connexity.staging.wpengine.com/wp-content/uploads/2016/01/CNX_AudienceModeling_9-15-15_email_vers.pdf

Constantinides, E., C.L. Romero, and M. A. Gómez Boria. 2009. Social Media: A New Frontier for Retailers? In *European Retail Research*, ed. Swoboda, B., D. Morschett, T. Rudolph, P. Schnedlitz, and H. Schramm-Klein, 1–28. European Retail Research. Wiesbaden: Gabler Verlag. https://doi.org/10.1007/978-3-8349-8099-1_1.

De Choudhury, M., M. Gamon, S. Counts, and E. Horvitz. 2013. Predicting Depression via Social Media. *AAAI Publications, Seventh International AAAI Conference on Weblogs and Social Media* 2013: 128–137.

Facebook. 2017. People-Based Marketing: Thinking People-First—Planning and Measurement. https://www.facebook.com/iq/articles/the-future-of-marketing-people-based-planning-and-measurement. Accessed 1 June 2019.

IBM. 2017. *Watson Personal Insights*. https://www.ibm.com/watson/developercloud/personality-insights/api/v3/. Accessed 27 Sept 2017.

Kosinski, M., D. Stillwell, and T. Graepel. 2013. Private Traits and Attributes are Predictable from Digital Records of Human Behavior. *PNAS* 110 (15): 5802–5805.

Kosinski, M., Y. Wang, H. Lakkaraju, and J. Leskovec. 2016. Mining Big Data to Extract Patterns and Predict Real-Life Outcomes. *Psychological Methods* 21 (4): 493–506.
Lambiotte, R., and M. Kosinski. 2014. Tracking the Digital Footprints of Personality. *Proceedings of the IEEE* 102 (12): 1934–1938.
LeCun, Y., Y. Bengio, and G. Hinton. 2015. Deep Learning. *Nature* 521: 436–444.
Lee, K., J. Mahmud, J. Chen, M. Zhou, J.Nichols. 2014. Who Will Retweet This? Automatically Identifying and Engaging Strangers on Twitter to Spread Information. *Proceedings of the 19th International Conference on Intelligent User Interfaces*, 247–256.
Leskovec, J., A. Rajaraman, and J.D. Ullman. 2014. *Mining of Massive Datasets.* Stanford textbook. http://infolab.stanford.edu/~ullman/mmds/book.pdf. Accessed 1 Sept 2017.
Li, W., X. Wang, R. Zhang, Y. Cui, J. Mao, and R. Jin. 2010. Exploitation and Exploration in a Performance Based Contextual Advertising System. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 27–36.
Quantcast. 2016. Understanding Digital Audience Measurements. https://www.quantcast.com/resources/understanding-digital-audience-measurement/. Accessed 1 June 2019.
Schwartz, H.A., J.C. Eichstaedt, M.L. Kern, L. Dziurzynski, S.M. Ramones, et al. 2013. Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLoS ONE* 8 (9): e73791.
Shmueli, G. 2017. Analyzing Behavioral Big Data: Methodological, Practical, Ethical, and Moral Issues. *Quality Engineering* 29 (1): 57–74.
Teich, C. 2016. Targeting in the Age of Micro-Segmentation: Good, Better and Best Approaches. http://www.hitwise.com/articles/targeting-age-microsegmentation/
Weed, K. 2017. The Future of Consumer Marketing? The CMO of Unilever Says It's 'Consumer Segments of One'. https://www.thinkwithgoogle.com/consumer-insights/unilever-consumer-marketing-segmentation/.
Yarkoni, T. 2010. Personality in 100,000 Words: A Large-Scale Analysis of Personality and Word Use Among Bloggers. *Journal in Research on Personality* 44 (3): 363–373.
Youyou, W., M. Kosinski, and D. Stillwell. 2015. Computer-Based Personality Judgments are More Accurate than Those Made by Humans. *PNAS* 112: 1036–1040.
Zeldes, Y., S. Theodorakis, E. Solodnik, A. Rotman, G. Chamiel, and D. Friedman. 2017. Deep Density Networks and Uncertainty in Recommender Systems arXiv:1711.02487.

References

Anderson, B. 2006. Imagined Communities. London: Verso [1983].
Anderson, C.W. 2011. Between Creative and Quantified Audiences: Web Metrics and Changing Patterns of Newswork in Local US Newsrooms. *Journalism* 12 (5): 550–566.
Andrejevic, M., and K. Gates. 2014. Editorial. Big Data Surveillance: Introduction. *Surveillance and Society* 12 (2): 185–196.
Andrejevic, M. 2016. Theorizing Drones and Droning Theory. In *Drones and Unmanned Aerial Systems: Legal and Social Implications for Security and Surveillance*, vol. 20, ed. A. Završnik, 43. New York: Springer.
Ang, I. 1991. *Desperately Seeking the Audience*. Abingdon: Routledge.
Bauman, Z., and D. Lyon. 2013. *Liquid Surveillance—A Conversation*. Cambridge: Polity.
Blackman, L. 2008. Affect, Relationality and the `Problem of Personality'. *Theory, Culture & Society* 25 (1): 23–47. https://doi.org/10.1177/0263276407085157.
Blackman, L. 2019. *Haunted Data: Affect, Transmedia, Weird Science*. London, New York: Bloomsbury Academic.
Blackman, L., J. Cromby, D. Hook, D. Papadopoulos, and V. Walkerdine. 2008. Creating Subjectivities. *Subjectivity* 22 (1): 1–27. https://doi.org/10.1057/sub.2008.8.
Boyd, D., and K. Crawford. 2012. Critical Questions for Big Data. *Information, Communication and Society* 15 (5): 662–679.
Brown, M. 2017. Understanding the 'Big Five' personality traits (OCEAN) and what they mean for marketers and brands. https://centiment.io/beta/blog/understanding-the-big-five-personality-traits-ocean-and-what-they-mean-for-marketers-and-brands/. Accessed 1 June 2019.
Cardon, D. 2015. *A Quoi rêvent les Algorithmes?*. Paris: Seuil.

Cheney-Lippold, J. 2011. A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture and Society* 28: 164–181.
Cheney-Lippold, J. 2017. *We are Data—Algorithms and the Making of Our Digital Selves*. New York: New York University Press.
Cover, R. 2006. Audience Inter/Active: Interactive Media, Narrative Control and Reconceiving Audience History. *New Media & Society* 8 (1): 139–58. https://doi.org/10.1177/1461444806059922.
Desrosières, A. 1988. Masses, Individus, Moyennes: la Statistique Sociale au XIXe siècle. C.N.R.S. Editions | « Hermès, La Revue » 1988/2 no. 2: 41–66.
Desrosières, A. 1998. *The Politics of Large Numbers—A History of Statistical Reasoning*. Cambridge: Harvard University Press.
Desrosières, A. 2008. *Pour une Sociologie Historique de la Quantification—l'Argument Statistique I*. Paris: Presses de l'Ecole des Mines.
Desrosières, A. 2014. *Prouver et Gouverner—Une Analyse Politique des Statistiques Publiques*. Paris: Editions de la Découverte.
Dumont, L. 1983. *Essai sur l'Individualisme—une Perspective Anthropologique sur l'Idéologie Moderne*. Paris: Seuil.
Ewald, F. 2011. Omnes and Singulatim. After Risk. *Carceral Notebooks* 7: 78–107.
Foucault, M. 1995. *Discipline and Punish*. New York: Vintage Books.
Foucault, M. 2009. *Security, Territory, Population. Lectures at the Collège de France. 1977-1978*, eds. M. Senellart, A.I. Davidson, F. Ewald and A. Fontana, trans. G. Burchell. Basingstoke: Palgrave Mcmillan.
Giraud, Eva. 2015. Subjectivity 2.0: Digital Technologies. *Participatory Media and Communicative Capitalism. Subjectivity* 8 (2): 124–146. https://doi.org/10.1057/sub.2015.5.
Goldberg, L.R. 1993. An Alternative 'Description of Personality': The Big-Five Factor Structure. *Journal of Personality and Social Psychology* 59 (6): 1216–29.
Goriunova, O. 2019a. Digital Subjects: An Introduction. *Subjectivity* 12 (1): 1–11. https://doi.org/10.1057/s41286-018-00065-2.
Goriunova, O. 2019b. Face Abstraction! Biometric Identities and Authentic Subjectivities in the Truth Practices of Data. *Subjectivity* 12 (1): 12–26. https://doi.org/10.1057/s41286-018-00066-1.
Habermas, J. 1991. *The Structural Transformation of the Public Sphere—An Inquiry into a Category of Bourgeois Society*. Cambridge: MIT Press.
Hartley, J. 2002. *Communication, Cultural and Media Studies: The Key Concepts*. London: Routledge.
Harvey, B. A Brief Personal History of Media Optimization. http://www.billharveyconsulting.com/articles/pdf/history-of-media-opt.pdf. Accessed 1 June 2019.
Karakayali, N., Burc K., and Idil G. 2018. Recommendation systems as technologies of the self: Algorithmic control and theformation of music taste. *Theory, Culture & Society* 35 (2): 3–24.
Kennedy, H., and G. Moss. 2015. Known or Knowing Publics? Social Media Data Mining and the Question of Public Agency. *Big Data & Society* 2 (2): 1–11. https://doi.org/10.1177/2053951715611145.
Koopman, C. 2019. *How We Became Our Data*. Chicago: The University of Chicago press.
Litt, E. 2012. Knock, Knock. Who's There? The Imagined Audience. *Journal of Broadcasting & Electronic Media* 56 (3): 330–45. https://doi.org/10.1080/08838151.2012.705195.
Livingstone, S. 2003. The Changing Nature of Audiences: From the Mass Audience to the Interactive Media User. In *Companion to Media Studies*, ed. A. Valdivia, 337–59. Oxford: Blackwell Publishing.
Livingstone, S. 2005. On the Relation Between Audiences and Publics. In: Livingstone, S., (ed.) *Audiences and Publics: When Cultural Engagement Matters for the Public Sphere*. Changing Media—Changing Europe series (2). Intellect Books, Bristol:17–41.
Marwick, A.E., and D. Boyd. 2011. I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society* 13 (1): 114–33.
Matthews, Julian. 2008. A Missing Link? The Imagined Audience, News Practices and the Production of Children's News. *Journalism Practice*: 264–279.
Mayer-Schönberger, V., and K. Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Kindle Android version.
Napoli, Philip M. 2010. Revisiting 'Mass Communication' and the 'Work' of the Audience in the New Media Environment. *Media, Culture and Society* 32 (3): 505–16.
Parisi, L. 2016. Automated Thinking and the Limits of Reason. *Cultural Studies <=> Critical Methodologies* 16 (5): 471–81.
Parisi, L. 2019. *The Alien Subject of AI. Subjectivity* 12 (1): 27–48.

Pasquale, F. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Kindle Android version.

Pentland, A. 2014. *Social Physics: How Social Networks Can Make Us Smarter.* Penguin, Kindle Edition.

Rouvroy, A. and T. Bern. 2013. Algorithmic Governmentality and the Prospects of Emancipation—Disparateness as a precondition for individuation through relationships?, *La Découverte « Réseaux »*, 2013/1 No 177: 163–196.

Scannell, Paddy. 2000. For-Anyone-as-Someone Structures. *Media, Culture and Society* 22 (1): 5–24.

Siegel, E. 2016. *Predictive Analytics—The Power to Predict Who will Click, Buy, Lie or Die*. Hoboken: Wiley.

Sternberg, Josh. 2013. "Retailers as Publishers." *Digiday* (blog). https://digiday.com/marketing/sociomantic-retailers-as-publishers/. Accessed 31 July 2013.

Tausczik, Y.L., and J.W. Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 29 (1): 24–54.

Turow, J. 2012. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth*. Kindle Android version.

Turow, J., and N. Draper. 2014. Industry Conceptions of Audience in the Digital Space. *Cultural Studies* 28 (4): 643–656.

Turow, J., L. McGuigan, and E.R. Maris. 2015. Making Data Mining a Natural Part of Life: Physical Retailing, Customer Surveillance and the 21st Century Social Imaginary. *European Journal of Cultural Studies* 18 (4–5): 464–478.

Zuboff, S. 2015. Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology* 2015 (30): 75–89.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Laurence Barry is an independent actuarial consultant and an associate lecturer at the Hebrew University of Jerusalem. She holds a PhD in Political Philosophy from the Hebrew University. Her thesis, entitled "Foucault and Postmodern Conceptions of Reason," deals with the intertwinement of knowledge and power in a historical perspective. Following Foucault's insights, she is currently working on the digital knowledge/power nexus.

Eran Fisher is a senior lecturer at the department of sociology, political science, and communication at the Open University of Israel. He received his PhD in sociology from the new school for social research in New York. He studies the intersection of digital technology and society. His books include Media and New Capitalism in the Digital Age (2010, Palgrave), Internet and Emotions (co-edited with Tova Benski, 2014, Routledge), and Reconsidering Value and Labour in the Digital Age (co-edited with Christian Fuchs, 2015, Palgrave).